Algorithmic composition of melodies with deep recurrent neural networks

Florian Colombo, Alexander Seeholzer, Samuel P. Muscinelli, Johanni Brea & Wulfram Gerstner

School of Life Sciences, Brain Mind Institute & School of Computer and Communication Sciences

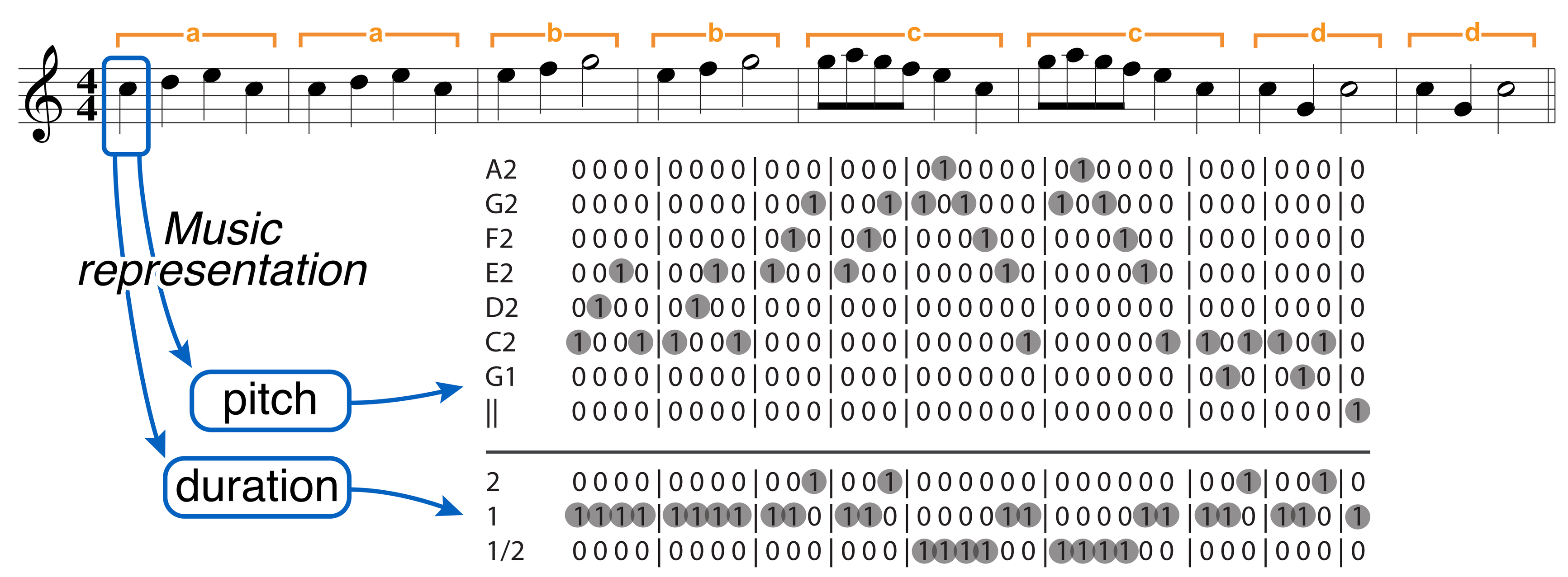
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Introduction

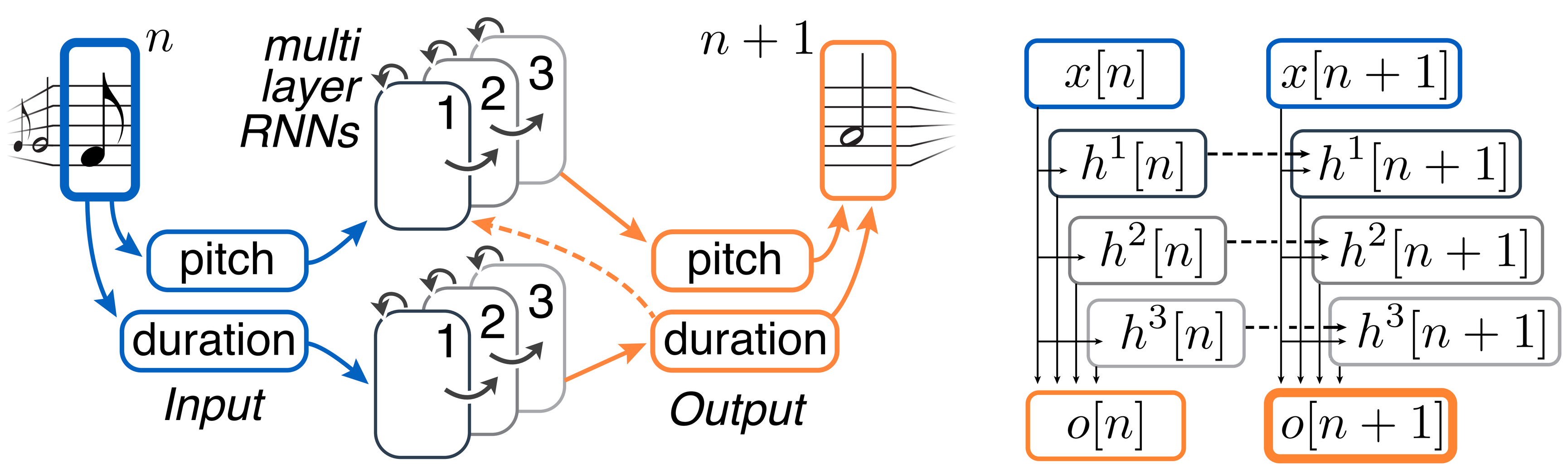
A big challenge in algorithmic composition is to devise a model that is both easily trainable and able to reproduce the **long-range temporal dependencies** typical of music. Here we investigate how artificial neural networks can be trained on a large corpus of melodies and turned into automated music composers able to generate new melodies coherent with the style they have been trained on.

We employ **gated-recurrent unit** (GRU) networks that have been shown to be particularly efficient in learning complex sequential activations with arbitrary long time lags [1]. Our model processes rhythm and melody in parallel while modeling the relation between these two properties. Using such an approach, we were able to generate interesting complete melodies or suggest possible continuations of a melody fragment that is coherent with the characteristics of the fragment itself.

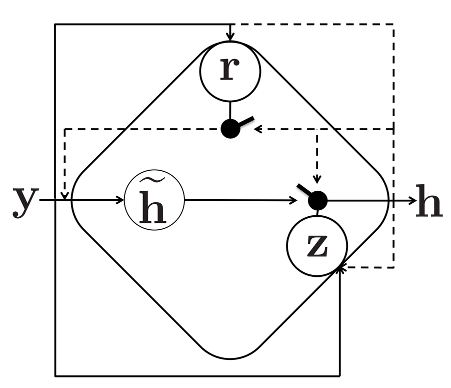
Music representation



Model



The hidden layers are composed of 128 GRUs



$$\mathbf{h}^i[n] = \mathbf{z}^i[n] \odot \mathbf{h}^i[n-1] + (\mathbf{1} - \mathbf{z}^i[n]) \odot \tilde{\mathbf{h}}^i[n]$$

$$\tilde{\mathbf{h}}^i[n] = \tanh\left(W_{y^i h^i}\mathbf{y}^i[n] + \mathbf{r}^i[n] \odot W_{h^i h^i}\mathbf{h}^i[n-1]\right)$$

$$\mathbf{z}^i[n] = \sigma\left(W_{y^i z^i}\mathbf{y}^i[n] + W_{h^i z^i}\mathbf{h}^i[n-1] + \mathbf{b}^i_z\right)$$

$$\mathbf{r}^i[n] = \sigma\left(W_{y^i r^i}\mathbf{y}^i[n] + W_{h^i r^i}\mathbf{h}^i[n-1] + \mathbf{b}^i_r\right)$$

$$\mathbf{o}_j[n] = \Theta\left(W_{y^o o}\mathbf{y}^o[n] + \mathbf{b}^o\right)_j = Pr\left(\mathbf{x}_j[n+1] = 1\right)$$

Training & Melody generation

- Minimize the log likelihood with Adam optimizer [2]

$$\mathcal{L}(\theta \mid \mathbf{x}^{1:S}) = \sum_{s=1}^{S}\sum_{n=1}^{N_s-1} \log\left(Pr\left(\mathbf{x}^s_j[n+1] = 1 \mid \mathbf{x}[n], \mathbf{h}^i[n-1], \theta\right)\right)$$

- Datasets
  - 2158 tunes from the Irish music corpus of Henrik Norbeck[a]
  - 1012 tunes from the Nottingham Music Database[b]
- Generate melodies $\tilde{\mathbf{x}}[n+1] \sim \mathrm{Cat}(\mathbf{o}[n])$

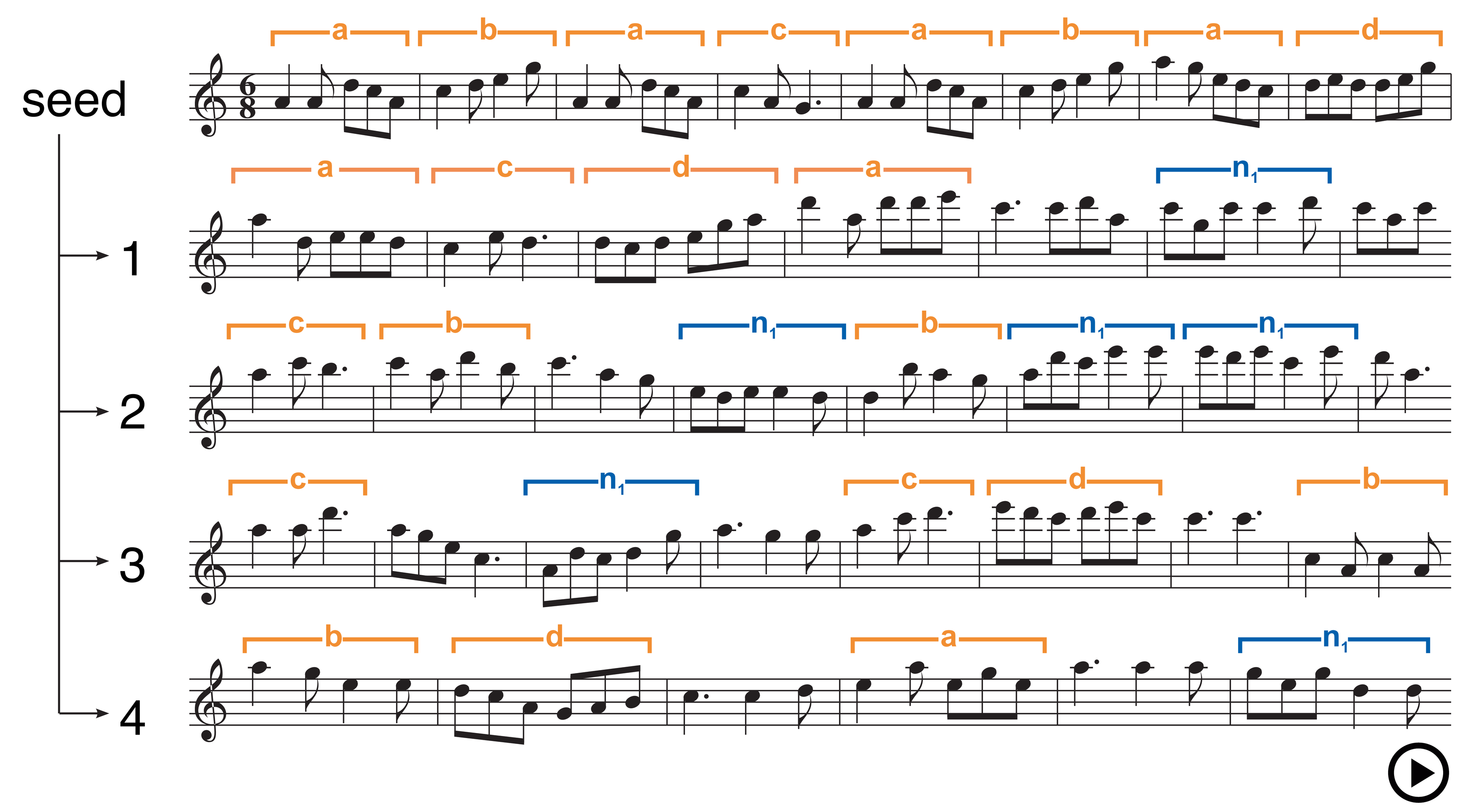
[a] http://www.norbeck.nu/abc/
[b] http://abc.sourceforge.net/NMD/

Acknowledgement

Research was supported by the Swiss National Science Foundation (200020_147200) and the European Research Council grant no. 268689 (MultiRules).

Song continuation

Four possible continuations suggested by the model



Autonomous song generation

An Irish tune autonomously generated by the model

Summary

- Symbolic music is represented by its melody and rhythm components
- We trained GRU-RNNs to **predict upcoming notes** in large corpora of music
- The intrinsic ability of GRU-RNNs to **operate on multiple timescales** enables them to learn complex temporal dependencies in music
- **Computer-aided algorithmic composition** is obtained by optimization of the model generalization ability but long-range temporal dependencies are missing
- Solutions that optimize the likelihood given the training data allow for **autonomous composition of complete and new melodies** exhibiting temporal dependencies on long timescales
- We showed the ability of GRU-RNNs coupled with the pitch/duration representation to compose interesting, complex and novel melodies

Future work

- Evaluation of generated melodies
- Hybridize GRU-RNN with other statistical models (e.g. Markov chain, HMM)
- Introduce harmonization and counterpoint
- Explore other (bigger) datasets – automated music parser

References

[1] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[2] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Contact Florian Colombo – phone: +41216931837 – email: florian.colombo@epfl.ch